



MODERN
COMMUNICATIONS

现代通信理论与技术丛书

Multimedia Communication

多媒体通信

● 吴 炜 编著

西安电子科技大学出版社
<http://www.xduph.com>

现代通信理论与技术丛书

多媒体通信

吴 炜 编著

西安电子科技大学出版社

2008

内 容 简 介

多媒体通信是计算机、通信和多媒体技术相结合的产物,目前已经成为通信的主要方式之一。多媒体通信主要研究多媒体数据的表示、存储、恢复和传输。

本书内容侧重于多媒体通信的关键技术,全书共分8章。第1章介绍相关概念;第2、3章介绍多媒体数据压缩编码技术和相关压缩编码标准;第4、5章论述视频通信中的编码速率控制技术和差错控制技术;第6章介绍实时多媒体业务传输协议;第7章描述多媒体通信终端;第8章讨论多媒体同步问题。

本书可作为高等院校相关专业的高年级本科生和研究生的教材或参考书,也可供从事多媒体通信、图像通信、通信工程技术和计算机应用技术等领域工作的科技人员参考。

图书在版编目(CIP)数据

多媒体通信/吴炜编著. —西安:西安电子科技大学出版社,2008.2

(现代通信理论与技术丛书)

ISBN 978-7-5606-1960-6

I. 多… II. 吴… III. 多媒体—计算机通信 IV. TN919.85

中国版本图书馆 CIP 数据核字(2007)第 204870 号

策 划 臧延新

责任编辑 阎 彬 臧延新

出版发行 西安电子科技大学出版社(西安市太白南路2号)

电 话 (029)88242885 88201467 邮 编 710071

http://www.xduph.com E-mail: xdupfxb@pub.xaonline.com

经 销 新华书店

印刷单位 陕西光大印务有限责任公司

版 次 2008年2月第1版 2008年2月第1次印刷

开 本 787毫米×1092毫米 1/16 印张 13

字 数 303千字

印 数 1~4000册

定 价 20.00元

ISBN 978-7-5606-1960-6/TN·0399

XDUP 2252001-1

*** 如有印装问题可调换 ***

本社图书封面为激光防伪覆膜,谨防盗版。

前 言

多媒体通信是近些年迅速发展起来的一门新兴的交叉性技术学科，它涉及通信、计算机和多媒体技术等领域，目前已经成为世界上发展最快和最富有活力的高新技术之一。多媒体通信的广泛应用可以极大地提高人们的工作效率，减轻社会的交通运输负担，对人们传统的教育和娱乐方式产生革命性的影响。

多媒体通信主要研究多媒体数据的表示、存储、恢复和传输。多媒体数据是由在内容上相互关联的文本、图像、图形、音频、视频和动画等多种媒体数据构成的一种复合信息实体。其中，有着严格时间关系的音频、视频等类型的数据称为连续媒体数据，其他类型的数据称为离散媒体数据。一般来说，多媒体数据至少包含两种媒体数据，其中一种必须为连续媒体数据。

多媒体通信的关键技术主要包括多媒体数据压缩编码、编码速率控制、差错控制、实时多媒体业务传输协议、多媒体通信终端和多媒体同步等。本书在介绍这些关键技术的基础上，全面阐述多媒体通信的基本原理和最新发展动态。

本书共分8章。第1章介绍多媒体和多媒体通信的相关概念，并描述传输多媒体数据的网络。第2章主要介绍多媒体数据压缩编码技术，包括率-失真理论、量化、变换编码、预测编码、熵编码、形状和纹理编码以及可伸缩视频编码等内容。第3章介绍多媒体数据压缩编码标准，包括静止图像、音频和视频的压缩编码标准等内容，并综述了图像和视频编码的发展趋势和未来展望。第4章介绍视频通信中的编码速率控制技术，包括速率控制技术的分类、速率控制的实现、现有视频压缩编码标准中的速率控制技术等内容，并综述了速率控制技术的研究进展。第5章介绍视频通信中的差错控制技术，包括误码发生的原因和现象，传输层差错控制、编解码端差错控制、解码端差错控制和交互式差错控制等四类差错控制技术，以及现有视频压缩编码标准中采用的差错控制技术等内容。第6章介绍实时多媒体业务传输协议，包括传输层协议和会话控制协议。第7章介绍多媒体通信终端。第8章介绍多媒体通信中的媒体同步，内容包括基本概念、参考模型、同步要求、同步描述的方法、分布式多媒体系统中的同步以及媒体同步控制技术，并提出了一种可行的媒体同步控制算法。

编写本书时，常义林教授在总体设计上提出了许多宝贵的建议，并对全书内容进行了详细认真的审阅，在此表示衷心的感谢。沙丽娜女士帮助绘制了许多插图，在此表示诚挚的谢意。

在本书的出版过程中，得到了西安电子科技大学出版社臧延新及其他老师的诸多帮助。借本书出版之际，向他们表示衷心的感谢。

在本书的编写过程中，参考和引用了许多前人的研究成果、著作、论文和网页等资料，这些资料在本书的参考文献中已尽量列出。但由于写作过程较长，同时有些通过网络查找

的资料没有详细的原始出处，可能会遗漏一些著作和论文的著录，在此表示歉意，并向所有文献的著作者致以由衷的敬意。

由于多媒体通信技术发展迅速，新概念、新知识和新方法层出不穷，限于作者的视野和学识水平，书中难免有错漏之处，恳请广大读者批评指正。

编 者

2007.11

目 录

第 1 章 绪论	1
1.1 多媒体	1
1.1.1 媒体的含义	1
1.1.2 多媒体的含义	2
1.1.3 多媒体技术的发展历史和现状	2
1.2 多媒体通信	5
1.2.1 多媒体通信的特征	5
1.2.2 多媒体通信的关键技术	6
1.2.3 多媒体通信的网络	7
第 2 章 多媒体数据压缩编码技术	14
2.1 概述	14
2.2 信源、信息量和熵	14
2.3 率—失真理论	16
2.3.1 信道和信道容量	16
2.3.2 率—失真函数	17
2.3.3 限失真信源编码定理	19
2.4 量化	20
2.4.1 均匀量化	20
2.4.2 最优量化	21
2.5 变换编码	22
2.5.1 离散余弦变换	22
2.5.2 小波变换	24
2.6 预测编码	29
2.6.1 语音的线性预测编码	29
2.6.2 图像的预测编码	31
2.7 熵编码	37
2.7.1 哈夫曼编码	37
2.7.2 算术编码	38
2.7.3 游程编码	40
2.8 形状和纹理编码	40
2.8.1 形状编码	40
2.8.2 纹理编码	43
2.9 可伸缩视频编码	45
2.9.1 空域可伸缩视频编码	45
2.9.2 时域可伸缩视频编码	46
2.9.3 质量可伸缩视频编码	48

2.9.4	频率可伸缩视频编码	49
2.9.5	精细颗粒度的可伸缩视频编码	49
2.9.6	基于小波变换的可伸缩视频编码	50
2.9.7	视频编码国际标准中的可伸缩视频编码方法	52
第3章	多媒体数据压缩编码标准	54
3.1	静止图像压缩编码标准	54
3.1.1	JPEG	54
3.1.2	JPEG 2000	56
3.2	视频压缩编码标准	59
3.2.1	H. 261、MPEG-1 和 MPEG-2	61
3.2.2	H. 263、H. 263+ 和 H. 263++	67
3.2.3	MPEG-4	73
3.2.4	H. 264	77
3.2.5	AVS	82
3.3	音频压缩编码标准	86
3.3.1	G. 7xx 系列	87
3.3.2	MPEG 伴音系列	88
3.3.3	AVS 音频	89
3.3.4	杜比数码系列	90
3.4	图像和视频编码的发展趋势和展望	91
3.4.1	图像和视频编码的基本策略	91
3.4.2	压缩效率——从像素到对象	92
3.4.3	满足网络和终端用户的需求	93
3.4.4	小结	94
第4章	视频通信中的编码速率控制技术	95
4.1	速率控制技术的分类	95
4.1.1	基于率-失真模型的速率控制	96
4.1.2	基于图像内容与人眼视觉特性的速率控制	97
4.1.3	无线信道中强健性速率控制	98
4.1.4	精细颗粒可伸缩编码的速率控制	99
4.2	速率控制的实现	100
4.2.1	目标比特分配	100
4.2.2	目标比特分配的实现	100
4.3	H. 263+ 中的速率控制技术	104
4.4	MPEG-2 中的速率控制技术	105
4.5	MPEG-4 中的速率控制技术	106
4.6	H. 264 中的速率控制技术	107
4.6.1	JVT-F086	108
4.6.2	JVT-G012	109
4.7	速率控制技术的研究进展	115
4.7.1	R-D 模型	115
4.7.2	质量受限的速率控制	116
4.7.3	从客观最优化到主观最优化	117

4.7.4 基于混合结构的最优化	119
第5章 视频通信中的差错控制技术	121
5.1 引言	121
5.2 传输层差错控制	122
5.2.1 FEC	122
5.2.2 ARQ	123
5.2.3 顽健性打包和复用	123
5.2.4 不等差错保护	124
5.3 编码端差错控制	124
5.3.1 差错隔离	124
5.3.2 顽健性二进制编码	125
5.3.3 容错预测	126
5.3.4 具有不等差错保护的分层编码	126
5.3.5 多描述编码	127
5.3.6 联合信源信道编码	129
5.4 解码端差错控制	133
5.4.1 差错检测	133
5.4.2 差错掩盖	134
5.5 交互式差错控制	134
5.5.1 编码参数自适应调整	134
5.5.2 基于反馈信息的参考图像选择	135
5.5.3 基于反馈的差错跟踪	135
5.5.4 无等待重传	135
5.6 H.263+和H.263++中的差错控制技术	135
5.7 MPEG-4中的差错控制技术	136
5.8 H.264中的差错控制技术	138
5.9 AVS-M中的差错控制技术	143
第6章 实时多媒体业务传输协议	146
6.1 引言	146
6.2 传输层协议	146
6.2.1 TCP	146
6.2.2 UDP	147
6.2.3 RTP	147
6.2.4 RTCP	149
6.3 会话控制协议	150
6.3.1 RTSP	150
6.3.2 SIP	151
第7章 多媒体通信终端	152
7.1 概述	152
7.2 H.323	153
7.2.1 H.323协议	153
7.2.2 H.323系统	155

7.3 SIP	158
7.3.1 下一代网络和软交换	158
7.3.2 SIP 系统	160
7.3.3 SIP 和 H.323 的比较	161
7.3.4 SIP 在 IMS 中的应用	162
7.4 H.324	164
7.4.1 H.324 协议	164
7.4.2 H.324 系统	165
7.5 H.320	166
7.5.1 H.320 协议	166
7.5.2 H.320 系统	167
7.6 H.321、H.310 和 H.322	167
7.6.1 H.321	167
7.6.2 H.310	168
7.6.3 H.322	169
第 8 章 多媒体同步	170
8.1 基本概念	170
8.1.1 多媒体数据内部的约束关系	170
8.1.2 逻辑数据单元	171
8.1.3 媒体内同步、媒体间同步和组同步	172
8.2 媒体同步的参考模型	173
8.2.1 媒体层	173
8.2.2 流层	174
8.2.3 对象层	174
8.2.4 描述层	174
8.3 同步描述	174
8.3.1 同步要求	175
8.3.2 同步描述的方法	177
8.4 分布式多媒体系统中的同步	180
8.4.1 分布式多媒体系统结构	180
8.4.2 同步描述数据的传送	181
8.4.3 影响媒体同步的因素	182
8.5 媒体同步控制技术	182
8.5.1 面向时间的播放控制算法	183
8.5.2 面向缓冲区的播放控制算法	184
8.6 MPEG-2 媒体同步控制算法	184
8.6.1 引言	184
8.6.2 节目流或传输流中的时间标签	185
8.6.3 MPEG-2 解码器中常见的同步机制	188
8.6.4 新的媒体同步控制算法	191
参考文献	196

第1章 绪 论

自20世纪80年代以来,随着电子技术和大规模集成电路技术的发展,计算机、通信和广播电视这三个互相独立的技术领域,相互融合、相互渗透,形成了一门崭新的技术,即多媒体技术。多媒体技术最直接、最简单的体现是配之以声卡、显卡的多媒体计算机。它一出现,即在世界范围内、在家庭教育和娱乐方面得到广泛的应用,并由此激发了小型激光视盘(VCD和DVD)以及数字电视和高清晰度电视(HDTV)的迅速发展。多媒体技术的应用与发展,又反过来进一步加速了这三个技术领域的融合,使多媒体通信成为未来通信技术发展的主要方向之一。

1.1 多 媒 体

1.1.1 媒体的含义

多媒体一词译自英文“multimedia”,它由 multiple 和 media 复合而成,核心词是媒体。媒体在计算机领域中有两种含义:一是指存储信息的实体,如磁盘、光盘、磁带、半导体存储器等,中文常译为媒质;二是指传递信息的载体,如数字、文字、声音、图形和图像等,中文常译为媒介。多媒体中的媒体是指后者。国际电报电话咨询委员会(Consultative Committee on International Telegraphs and Telephones, CCITT)曾对媒体作过如下五种分类:

(1) 感觉媒体(perception medium)。感觉媒体是指能直接作用于人们的感觉器官,从而使人直接产生感觉的一类媒体,如声音、音乐、语音、文本、图形、静止和活动图像等。

(2) 表示媒体(representation medium)。表示媒体指为了加工、处理和传输感觉媒体而人为研究、构造出来的一种媒体。其目的是更有效地将感觉媒体从一地向另外一地传送,便于加工和处理。语言编码、文本编码和图像编码等都属于表示媒体。

(3) 表现媒体(presentation medium)。表现媒体是指感觉媒体和用于通信的电信号之间转换用的一类媒体。它又分为两种:一种是输入表现媒体,如键盘、摄像机和麦克风等;另一种是输出表现媒体,如显示器、打印机等。

(4) 存储媒体(storage medium)。存储媒体用于存放表示媒体,以便计算机能随时处理、加工和调用信息编码。这类媒体有纸张、磁带、磁盘、光盘等。

(5) 传输媒体(transmission medium)。传输媒体是用来将媒体从一处传送到另一处的物理载体,如双绞线、同轴电缆、光纤等。

可以看出,多媒体中的媒体一般指的是感觉媒体。

1.1.2 多媒体的含义

所谓“多媒体”，从文字理解就是“多种媒体的综合”，相关的技术也就是“怎样进行多种媒体综合的技术”。多媒体的定义或说法多种多样，人们从不同的角度出发，对多媒体进行了不同的描述。正因为如此，很多人往往会提出以下问题：电视算不算多媒体？可视图文呢？各种家电的组合呢？各种彩色画报呢？为什么以前也有计算机图形、图像，而不称之为多媒体呢？事实上人们普遍认为，“多媒体”是指能够同时获取、处理、编辑、存储和展示两个以上不同类型信息媒体的技术，这些信息媒体包括文字、声音、图形、图像、动画与视频等。因此，我们常说的“多媒体”最终被归结为一种“技术”，不是指多种媒体本身，而主要是指处理和应用它的一整套技术。因此，“多媒体”实际上常被当做“多媒体技术”的同义语。

现在的多媒体技术往往与计算机联系起来。这是由于计算机的数字化及交互式处理能力极大地推动了多媒体技术的发展。通常，可把多媒体看做是先进的计算机技术与视频、音频和通信等技术融为一体而形成的新技术或新产品。因此，我们为多媒体技术下的定义是：计算机综合处理文本、图形、图像、音频与视频等多种媒体信息，使多种信息建立逻辑连接，集成为一个系统并具有交互性。简而言之，多媒体技术就是计算机综合处理声、文、图、像等信息的技术，具有集成性、实时性和交互性。

根据多媒体技术的定义，我们可以看到它有三个显著的特点，即集成性、实时性和交互性，这也是它区别于传统计算机系统的特征。所谓集成性，一方面是指媒体信息即声音、文字、图像、视频等的集成，另一方面是指显示或表现媒体设备的集成，即多媒体系统一般不仅包括计算机本身，而且还包括电视、音响、录像机和激光唱机等设备。所谓实时性，是指在多媒体系统中声音和视频是实时的，多媒体系统提供了对这些时基媒体实时处理的能力。交互性是多媒体计算机与其他家用声像电器，像电视机、激光唱机等有所差别的关键特征。普通家用声像电器没有交互性，即用户只能被动收看，而不能介入到媒体的加工和处理之中。

所以，我们可以判断什么是“多媒体”。因为电视不具备像计算机一样的交互性，不能对内容进行控制和处理，它就不是“多媒体”；同理，各种家电的组合、画报也不是；仅有个别种类媒体的计算机系统也不是。而那些采用计算机集成处理多种媒体（一般包括声音、图像、视频和文字等）的系统，如多媒体咨询台、交互式电视、交互式视频游戏、计算机支持的多媒体会议系统、多媒体课件及展示系统等，都属于“多媒体”的范畴。

1.1.3 多媒体技术的发展历史和现状

几十年前，人们曾经把几张幻灯片配上同步的声音，称为多媒体系统。而今天，随着微电子、计算机和通信技术的发展，多媒体技术被赋予了新的内容，多媒体系统也发生了质的变化。

1. 启蒙发展阶段(1984—1990年)

1984年美国Apple公司在Macintosh上为了改善人机之间的接口，引入了位映射(Bitmap)的概念对图进行处理，并使用了窗口(Window)和图符(Icon)作为用户接口。该公司最早采用图形用户接口(GUI)取代了字符用户接口(CUI)，用鼠标器和菜单取代了键盘

操作。

美国 Commodore 个人计算机公司于 1985 年率先推出世界上第一台多媒体计算机 Amiga, 后来不断完善, 形成了一个完整的多媒体计算机系列。该公司的 Amiga 系列分别配置了 Motorola 公司生产的 M68000 微处理器系列, 并采用了自己研制的三个专用芯片 Agnus(8370)、Paula(8364)和 Denise(8362)。为了适应各类不同用户对多媒体技术的需要, Commodore 公司还提供了一个多任务 Amiga 操作系统, 它具有下拉菜单、多窗口、图符显示等功能。

1986 年, 荷兰 Philips 公司和日本 Sony 公司联合研制并推出了交互式紧凑光盘系统 (Compact Disc Interactive, CD-I), 同时还公布了 CD-ROM 文件格式, 得到了同行的认可并成为 ISO 国际标准。该系统把高质量的声音、文字、图形、图像都进行了数字化, 并将其放在 650 MB 的只读光盘上, 用户可以连到电视机上显示。后来, CD-I 随着 Motorola 公司微处理器的发展不断改进, 并广泛应用于教育、培训和娱乐领域。

早在 1983 年, 美国无线电公司 RCA 的研究中心 (David Sanaoff Research Center in Princeton, New Jersey) 就开始研究和开发相关系统设备。它以计算机技术为基础, 用标准光盘来存储和检索静止图像、活动图像、声音和其他数据。后来, RCA 把推出的交互式数字视频系统 (DVI) 卖给了美国通用电气公司。1987 年, Intel 公司收购了 DVI, 并经进一步的研究和改善后, 于 1989 年初把 DVI 技术开发成为一种可以普及的商品。此后, Intel 和 IBM 合作, 在 Comdex/Fall'89 展示会上展出了 Action Media 750 多媒体开发平台。当时 Action Media 750 的硬件由三块专用插板组成, 即音频板、视频板和多功能板。

1991 年, Intel 公司和 IBM 公司又推出多媒体改进技术 Action Media II, 可作为微通道和 ISA 总线的选件。它由两块板, 即采集板和用户板组成。DVI 软件开发出了多媒体的音频和视频内核 (Audio Video Kernel, AVK), 同时还开发了在 Windows 3.0 和 OS/2 1.3 下运行的 AVK。AVK 提供低层编程接口 Beta DV-MCI, 后来又扩展到 Windows 和 OS/2 上。世界上先后有几家公司为其开发软件, 在美国已经广泛使用。

2. 标准化和应用阶段(20 世纪 90 年代以来)

20 世纪 90 年代以来, 多媒体技术逐渐成熟, 从以研究开发为重心转移到以应用为重心。多媒体应用迅猛发展, 应用范围包括培训、教育、商业和产品展示、产品和事务咨询、信息出版、销售演示、家庭教育和个人娱乐等众多领域。

1993 年, PC 机在美国作为圣诞节礼物空前流行, 多媒体 PC 机引起购物者强烈的兴趣。市场上的多媒体产品使人目不暇接。美国市场上第二大多媒体供应商 Media Vision 公司总经理 Allan C. Thygeson 说: “1993 年是多媒体圣诞节, PC 机真正变成了消费品。”

由于多媒体技术是一种综合性技术, 其产品的应用目标既涉及研究人员也面向普通消费者, 因此标准化问题是多媒体技术实用化的关键。标准的出现推动了相关产业产值的大幅度增长, 产品成本和价格大幅度降低, 并大大改善了多媒体产品之间的兼容性, 导致产品应用的迅速增长。

1990 年 10 月, 在微软公司多媒体开发工作者会议上提出了多媒体 PC 机标准 MPC 1.0, 后来 MPC 理事会重新精练了多媒体 PC 机的定义, 去掉了 80286 处理器, 认为最低要用主频为 20 MHz 的 386SX。1993 年, 多媒体计算机市场委员会 (MPMC) 发布了多媒体个人计算机的性能标准 MPC 2.0。1995 年 6 月, MPMC 又宣布了新的多媒体个人计算机

技术规范 MPC 3.0。事实上,随着应用要求的提高和多媒体技术的不断改进,多媒体功能已成为新型个人计算机的基本功能。这样,就没有必要继续发布 MPC 的新标准了。

多媒体技术应用的关键问题是对多媒体数据进行压缩编码和解码。国际标准化组织(International Organization for Standardization, ISO)和国际电报电话咨询委员会(CCITT)成立了联合图像专家组(Joint Photographic Experts Group, JPEG),于1991年制定了第一个图像标准 ISO/IEC 10918,即“多灰度静止图像的数字压缩编码”,称为 JPEG 标准。这里的 IEC 是指国际电工技术委员会(International Electro-technical Commission)。JPEG 专家组于1999年制定的第二个标准是 JPEG-LS(ISO/IEC 14495),用于静止图像无损编码,并于2000年底制定了最新的静止图像压缩标准 JPEG 2000(ISO/IEC 15444)。

ISO 和 IEC 的共同委员会中的活动图像专家组(Moving Picture Experts Group, MPEG)于1992年制定的 MPEG-1(ISO/IEC 11172)标准,是为传输速率为 1.5 Mb/s 的数字声像信息的存储而制定的。MPEG-2(ISO/IEC 13818)是由活动图像专家组和 CCITT 于1992年改组的国际电信联盟电报电话部(International Telecommunications Union for Telegraphs and Telephones Sector, ITU-T)的第15研究组于1994年共同制定的。MPEG-2 是一个通用的标准,它克服并解决了 MPEG-1 不能满足日益增长的多媒体技术、数字电视技术、多媒体分辨率和传输率等方面的技术要求上的缺陷,能在很宽范围内对不同分辨率和不同输出比特率的图像信号有效地进行编码,编码速率为 4~100 Mb/s。活动专家组于1999年2月正式公布了 MPEG-4 V1.0 版本,同年12月又公布了 MPEG-4 V2.0 版本。MPEG-4 标准主要针对可视电话、视频电子邮件和电子新闻等应用,其传输码率要求较低,在 4800~6400 b/s 之间。MPEG 系列的其他标准还有 MPEG-7 和 MPEG-21。

H. 261 是 ITU-T 第15研究组于1984—1989年制定的针对可视电话和视频会议等业务的视频压缩标准,目的是在窄带综合业务数字网(N-ISDN)上实现速率为 $p \times 64$ kb/s 的双向声像业务,其中 $p=1\sim 30$ 。因此,H. 261 又称为 $p \times 64$ 标准。ITU-T 于1995年制定的甚低比特率视频压缩编码标准 H. 263,其传输码率可以低于 64 kb/s,该标准特别适用于无线网络、PSTN 和因特网等环境下的视频传输,所有的应用都要求视频编码器输出的码流在网络上进行实时传输。为了提高编码效率,增强编码功能,ITU-T 对 H. 263 进行了多次补充,补充修订的版本有1998年制定的 H. 263+ 及2000年制定的 H. 263++。H. 264 标准是由 ITU-T 的视频编码专家组(VCEG)和 ISO/IEC 的活动图像专家组共同成立的联合视频小组(Joint Video Team, JVT)于2003年3月发布的,也称为 MPEG-4 的第10部分,即高级视频编码(Advanced Video Coding, AVC)。H. 264 继承了以往标准的优点,并进一步提高了编码算法的压缩效率和图像播放质量。

除了国外制定的以上视频编码标准外,我国首次自主制定、具有自主知识产权的数字音/视频编/解码标准 AVS,是数字电视、IPTV 等音/视频系统的基础性标准。AVS 标准第2部分即视频部分属于高效的第二代视频编/解码技术,与 MPEG-2 相比,其编码效率提高 2~3 倍,且实现方案简洁。同时,AVS 标准具有专利许可方式简洁、相关标准配套的优势,这将为我国的 IPTV、数字电视广播等重大信息产业应用及民族 IT 产业发展起到积极的推动作用。

对于音频压缩, MPEG-1、MPEG-2、MPEG-4 和 AVS 等标准中都含有相应的音频压缩部分, 而且 ITU 还制定了一系列的音频压缩标准, 如 G. 711、G. 721、G. 722、G. 723.1、G. 728 和 G. 729 等。

多媒体技术应用的水平取决于市场的需求和多媒体技术的成功与否, 多媒体产品是否普及取决于其性价比和能否开拓更广泛的应用领域。多媒体技术的典型应用包括以下一些方面: 教育和培训、销售和演示、娱乐和游戏、管理信息系统、多媒体服务器、视频会议和计算机支持协同工作等。

1.2 多媒体通信

多媒体通信是计算机、通信和多媒体技术相结合的产物, 目前它已经成为通信的主要方式之一。多媒体通信的广泛应用将会极大地提高人们的工作效率, 减轻社会的交通运输负担, 对人们传统的教育和娱乐方式产生革命性的影响。

多媒体通信主要研究多媒体数据的表示、存储、恢复和传输。多媒体数据是由在内容上相互关联的文本、图像、图形、音频、视频和动画等多种媒体数据构成的一种复合信息实体。其中, 有着严格时间关系的音频、视频等类型的数据称为连续媒体(Continuous Media)数据, 其他类型的数据称为离散媒体(Discrete Media)数据。一般来讲, 多媒体数据至少包含两种媒体数据, 其中一种必须为连续媒体数据。

1.2.1 多媒体通信的特征

多媒体通信系统具有以下三种特征:

- 集成性, 是指多媒体通信系统能够对至少两种媒体数据进行处理, 并且可以输出至少两种媒体数据。

- 交互性, 是指多媒体通信系统中用户与系统之间的相互控制能力, 它包括两个方面: 其一是人与终端之间的交互, 即用户在使用多媒体通信终端时, 终端向用户提供的操作界面; 其二是终端和系统之间的通信。

- 同步性, 是指多媒体通信终端在显示多媒体数据时, 必须以同步方式进行, 这样就将构成的一个完整的信息显示在用户面前。

这三个特征是构成多媒体通信系统的基础, 缺一不可, 如果不能同时具备这些特征, 特别是交互性和同步性, 就不能称其为多媒体通信系统。在多媒体通信终端的用户对通信的全过程有着完备的交互控制能力。交互式是多媒体通信系统的一个主要特征, 是区别多媒体通信系统和非多媒体通信系统的一个主要准则, 如电视机不能提供用户对整个过程进行有效的选择控制, 而视频点播(Video On Demand, VOD)则可以, 所以说 VOD 系统是多媒体通信系统。同步性是多媒体通信系统最主要的特征之一, 信息的同步与否可以决定系统是多媒体系统还是多种媒体系统; 同时, 同步性也是多媒体通信系统中最为困难的技术问题之一, 这是因为多媒体通信系统是一个资源受限的系统, 它的通信速率和终端内存都受到限制。

1.2.2 多媒体通信的关键技术

多媒体通信的关键技术取决于多媒体通信的特征和多媒体信息的特点(数据量大且类型多,实时性要求高,应用领域广泛),需要解决的问题集中于压缩数据量、网络传输速度、快速存储和检索等。其关键技术主要包括多媒体数据压缩编码、编码速率控制、差错控制、实时多媒体业务传输协议、多媒体通信终端和多媒体同步等。

1. 多媒体数据压缩编码

人们的信息交换方式不局限于文字和数据,而更多的是通过多种器官来接收外界的信息,其中80%以上是通过视觉接收的。早期的多媒体通信系统,曾按照模拟方式,仿效家电技术中的磁带录音和磁带录像。虽然利用模拟设备把多媒体数据集中在一个原型系统中是可以接受的,但模拟方式所固有的缺点限制了多媒体通信的进一步发展,因此,在多媒体通信中主要采用数字化技术。在大量的多媒体信息的数字化过程中,最初的图形和声音数字化后的代码量是可以接受的,但是后来在对音频和活动图像信息进行数字化后,产生的代码量与目前的硬件存储技术和硬件传输技术产生了矛盾,这是因为目前的硬件技术的发展远远不能满足大量的数字化信息的存储和传输需求,这一矛盾已经成为制约多媒体通信发展的一个瓶颈问题。显然,要真正把多媒体通信投入应用,首先必须解决的是多媒体数据的压缩编码和解码。只有有效地对大量的多媒体数据进行压缩,多媒体通信才能得到很好的发展和应用。

2. 编码速率控制

在多媒体通信中,音频业务和视频业务要求实时传输,对时延十分敏感,不过可以容忍一定程度的误码。相对于音频业务,视频业务对网络带宽有较高的需求。由于视频编码过程的固有特性(变换、量化和熵编码等)以及视频内容的变化,编码后的视频数据是变比特率的,这为其在网络上传输带来了很大困难。为了在信道带宽和传输时延有限的情况下高效传输视频数据,使得网络视频业务具有一定的播放质量,通常需要对视频编码进行速率控制;而对于带宽不受限的存储视频来说,由于解码端的缓冲区是有限的,因此仍然需要对编码视频的速率进行控制。只有对视频码流进行很好的速率控制,才能满足给定的带宽条件,改善视频编码效果。

3. 差错控制

在目前的网络环境中,高效压缩后的视频码流存在很强的相关性,对传输误码非常敏感。现有的视频压缩编码标准主要采用预测编码和变长编码技术,一旦网络传输发生误码,往往使得解码器失去同步,而后续码流即使可以正确接收,也不能正确解码,造成误码扩散,最终使视频重建质量严重下降。为了保证视频码流在网络中能可靠传输,必须根据视频业务的特性和网络的特点采用差错控制技术。

4. 实时多媒体业务传输协议

实时多媒体业务传输协议包括传输层协议和会话控制协议。传输层协议为流媒体应用提供端到端的网络传输功能,会话控制协议通过在一个会话期间定义消息和进度来控制多媒体数据的传送。这些协议保证了多媒体业务能够在网络上进行实时通信。

5. 多媒体通信终端

多媒体通信系统的集成特性对标准化提出了很高的要求。多媒体通信终端是能够集成多种媒体数据,通过同步机制将多媒体数据呈现给用户,具有交互功能的新型通信终端,是多媒体通信系统的重要组成部分。随着多媒体通信技术的发展,已经开发出一系列多媒体通信终端的相关标准和设备,它们又反过来促进多媒体通信的发展。

6. 多媒体同步

由于多媒体数据至少包含两种媒体数据,其中一种必须为连续媒体数据,而连续媒体数据有着严格的时间关系,因此存在多媒体同步问题。可以说,多媒体同步是由多媒体数据所具有的独特特征而引发的问题;也可以说,只有在多媒体通信系统中才有多媒体同步问题。例如在远程教学中,幻灯片必须与其解说词保持同步,否则屏幕上正在播映的幻灯片和观众听到的讲解人的解说会不一致。作为多媒体通信中的一项关键技术,多媒体同步已经引起了学术界的广泛关注。

1.2.3 多媒体通信的网络

多媒体通信具有集成、交互和同步的特征,传输的多媒体数据量大、实时性强、对时延抖动敏感,所以多媒体通信对网络有一定的要求。本书从多媒体通信的角度出发,主要介绍多媒体通信中的关键技术,而不会单独分章来描述传输多媒体数据的网络,所以本章花少量篇幅向读者概括介绍一下目前的网络。

在介绍网络之前,先了解一下网络的分层结构。为了降低网络设计的复杂度,大多数网络是按照一定的分层结构来组织的,其中每一层建立在下层的基础之上。每一层都有协议,这里协议是指通信双方关于通信过程如何进行所达成的共同约定,这些约定规定了计算机交换的消息的语法和含义。每一层都是为它的上层服务的,逻辑上相关的服务被划分为单独的一层且定义在参考模型中。OSI参考模型就是其中的一种参考模型,如图1.1所示。

由于OSI参考模型用一种简单的方法对网络中的复杂硬件和协议软件进行了清晰的描述,因此得到广泛的应用。在此模型中,最低层对应于硬件,其他各层分别对应于使用该硬件的固件和软件。下面是每层的简单描述。

- 物理层(physical layer): 对应于基本的网络硬件,涉及到通信在信道上传输的原始比特流。设计上必须保证一方发出二进制“1”时,另一方准确无误地收到。这里的设计主要是处理机械的、电气的和过程的接口,以及物理层下的物理传输介质等问题。

- 数据链路层(data link layer): 主要任务是加强物理层传输原始比特的功能,使之对网络层显现为一条无错线路。该层规定了数据按照什么样的格式被组织成帧以及如何在网络中传输数据帧。目前有两种类型的网络,即采用点到点连接的网络和采用广播信道的网

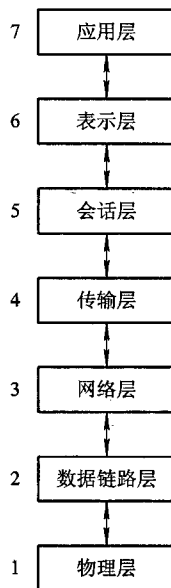


图 1.1 OSI 参考模型

络。在广播式网络中有一个新问题，即如何控制对共享信道的访问。数据链路层中有一个特殊的子层，即介质访问控制(Media Access Control, MAC)层，就是用来解决这个问题的。

- 网络层(network layer): 主要用来控制子网的运行，其中最关键的问题是分组在信源到信宿之间如何选择路由。该层规定了如何分配地址以及如何将分组从信源传送到信宿。寻址、交换、路由、拥塞控制和分组排序等都是网络层的任务。

- 传输层(transport layer): 主要任务是从端到端有效、可靠地通过网络透明地传送报文。该层根据通信子网的特性，最佳地利用网络资源，并以可靠和经济的方式为两端的会话层提供建立、维护和取消传输连接的功能，可靠地传输数据。

- 会话层(session layer): 允许不同主机上的用户建立会话关系。该层可以像传输层那样传送普通数据，而且还提供一种增强型的服务。会话层主要规定了如何与远程系统建立一个通信的会话，如登录到远程的分时系统。

- 表示层(presentation layer): 完成某些特定的功能，通过这些功能为用户提供一种通用的解决方法，而不是让每个用户自己去完成。表示层以下各层只关注如何可靠地传送数据，而表示层关注的是传送数据的语义和语法。

- 应用层(application layer): 提供与用户应用进程的接口。该层包含很多被广泛使用的协议，通过使用虚拟终端软件来解决终端的兼容问题。它还提供文件传输的功能。

后来，人们把会话层、表示层和应用层合为一个应用层，我们在本书中将遵循这种定义。

对于网络的分类方法，有两种因素是非常重要的：传输技术和距离尺度。一般来说，目前普遍使用的传输技术有两种，分别为广播式链接和点到点链接。广播网络只有一个通信信道，网络上所有的主机都共享该信道。在广播系统中，若一个发送分组被网络中每一台主机接收并进行处理，则这种模式称为广播；若该分组被网络中一组主机即网络中所有主机的一个子集接收并处理，则该模式称为多播。点到点网络是由许多连接构成的，每一个连接对应于一对主机。这种只有一个发送方和一个接收方的点到点传输模式称为单播。

网络分类的另一个准则是网络的距离尺度。根据终端之间的距离，网络可以分为以下几类：个人区域网络(Personal Area Network, PAN)、局域网(Local Area Network, LAN)、城域网(Metropolitan Area Network, MAN)、广域网(Wide Area Network, WAN)和互联网。个人区域网络仅仅供一个人使用。如一台计算机通过一个无线网络和它的附属设备连接起来，这个无线网络就是个人区域网络。局域网是专有网络，通常指一个建筑物内或者一个校园内的独立的计算机网络。它的特征在于所能覆盖的范围以及较高的传输速率，它可以覆盖几千米的范围，连接有限的主机数量。为了容纳更多的主机，一些局域网可以彼此连接起来，组成一个更大规模的网络——城域网。城域网所采用的技术基本上和局域网类似，只是规模上更大一些，通常覆盖市内的区域。一个城域网一般属于一个公共或私有的网络运营商，受到运营商的管理，可以连接数以千计的主机。广域网又称为远程网，跨越了一个很大的范围，通常是一个国家或者一个洲。广域网包含了大量的主机，在这些主机上可以运行应用程序，并且这些主机是通过通信子网连接起来的。两个或者多个广域网相互连接起来之后称为互联网。

传输多媒体数据的局域网类型有以太网(Ethernet)、令牌环网、100VG AnyLAN、光